Conditional Entropy for Deception Analysis

John Custy*
SPAWAR Systems Center, Code 24527
San Diego, CA 92152
john.custy@navy.mil

Neil C. Rowe
Naval Postgraduate School
Department of Computer Science
Monterey, CA 93940
ncrowe@nps.edu

***Abstract*—This paper describes how basic concepts from information theory can be used to analyze deception. A simple channel model known as a Z-channel provides a fundamental description of deception, and allows deception performance to be measured as the conditional entropy, or equivocation, of the channel. The Z-channel model suggests that associated with any given deception is another deception we call the "symmetric complement" of the given deception. A deception deployed along with its symmetric complement can provide implementation and/or performance benefits to a deceiver, but can also open counterdeception opportunities for the deception target. Finally, two deception-based computer security techniques are described and analyzed. A *fake honeypot* can be used to inoculate a computer against intrusions, and *spoofing channels* provide means to safely and effectively respond to computer intrusions. The spoofing channel deception is of fundamental interest because it is equal to its symmetric complement.**

***Index Terms*—deception, computer security, information theory, spoofing channels**

I. Introduction

One of the most powerful ways of countering an adversary is through deception. Intuition suggests that deception and communication are in a sense "opposites" or "duals," but any relationship that may exist has never been made precise. This is unfortunate because sophisticated and useful mathematical tools have been developed to characterize communication systems [1], yet few if any techniques currently exist for the mathematical analysis of deception [15].

In this paper we use elementary concepts from information theory to model deception and to characterize deception performance. In addition, we present two ideas for computer security that arose from our study of deception and information theory. Our work is based on the idea that deception is the act of manipulating observations made by a deception target so as to imply a specific false version of reality. This allows the average performance of a deception to be evaluated as if it were a communication channel, with the deceivers actions playing the role of noise.

It is important to note that our approach gives information only about the average effectiveness of a deception, and not about the outcome of any particular deception encounter. In a similar way, analysis of a communication system will usually not reveal whether a specific symbol transmitted at a specific instant will be received correctly. Rather, in both cases performance is characterized only in terms of averages. Our model is thus most useful for characterizing deceptions that are applied repeatedly, such as spam, phishing ploys, and computer security methods such as those presented in this paper.

Our paper is structured as follows. Section II establishes terminology and introduces some of the main concepts used in the remainder of the paper. Section III gives a definition for deception, and shows how the definition conforms to a channel model known as a Z-channel. Examples illustrate how common deceptions can be mapped to the Z-channel model. Section IV shows how the mathematical tools used to describe communication systems, in particular conditional entropy, can be used with the Z-channel model to quantify deception performance.

The Z-channel model of deception suggests that associated with a given deception is a closely related deception we refer to as a symmetric complement. The link between a deception and its symmetric complement is described and illustrated in Section V. Finally, in Section VI we describe two software tools, the fake honeypot and the spoofing channel, that use deception to support computer security.

II. A Sketch of Communication Terminology & Concepts

Our discussion of deception will rely heavily on abstract representations of binary communication systems like those in Figures 1 and 2. In informal terms, these figures describe two different ways in which random errors can be introduced into a stream of binary symbols. Figure 1 shows a *Z-channel*, which can introduce only one type of error: the transmitted symbol B will be received as $\tilde{A}$ at a rate of p_{BA} and as $\tilde{B}$ at a rate of $1 - p_{BA}$, but the symbol A will always be received as $\tilde{A}$. Figure 2 shows what we will call a *Binary Symmetric Channel*, or BSC, which can introduce two types of errors: a transmitted symbol A can be received as either $\tilde{A}$ or $\tilde{B}$, and a B can be received as either $\tilde{A}$ or $\tilde{B}$. Clearly, the Z-channel is just a special case of the BSC. (In common usage the term "Binary Symmetric Channel" refers only to the case where $p_{AB} = p_{BA}$, while we will use the term to refer to the case in which both p_{AB} and p_{BA} are non-zero, but are not necessarily equal.)

Channel models like those in Figures 1 and 2 are fully characterized by the transition probabilities and by the *a priori* input probabilities p_A and $p_B = 1 - p_A$, and by the transition

*Supported in part by the SPAWAR Systems Center In-house Laboratory Independent Research Program.

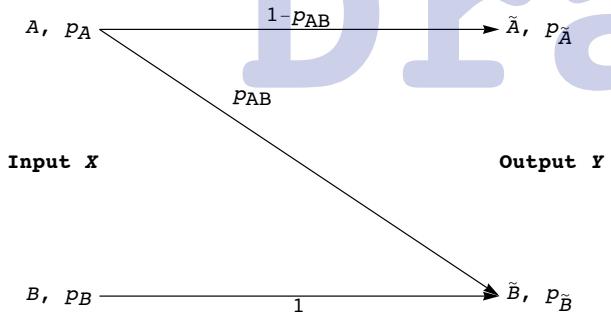

Fig. 1. An abstract representation of a discrete binary communication channel known as a Z-channel. This is a special case of the channel shown in Figure 2.

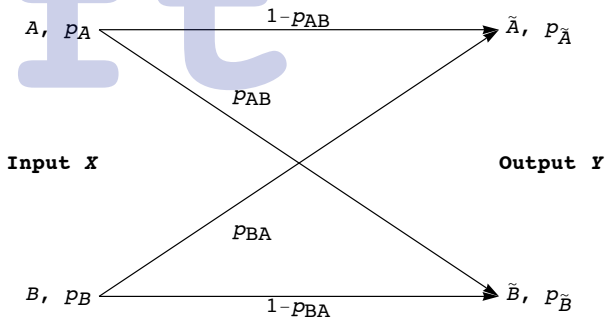

Fig. 2. An abstract representation of a discrete binary communication system. We refer to this as a Binary Symmetric Channel, even when $p_{AB} \neq p_{BA}$.

probabilities p_{AB} and p_{BA}. However, these probabilities can be cast into a different and extremely useful form. This alternate form provides essentially a *count* of the number of sequences that satisfy a given set of probabilistic constraints. As a rough introduction, consider the number of sequences of A's and B's consisting of a single B and nine A's: there are $\binom{10}{1} = 10$ such sequences. Similarly, there are $\binom{10}{2} = 45$ sequences consisting of two B's within a sequence of eight A's, and there are $\binom{10}{5} = 252$ sequences consisting of five B's and five A's. As the number of B's increases further, the number of possible sequences decreases.

Of course, there is nothing significant about sequences of length ten, and counts of this sort can be carried out for sequences of any length. Thus the number of ordered sequences of length N consisting of pN copies of the symbol A's and $(1-p)N$ copies of the symbol B's is $\binom{N}{pN}$. It turns out that when N becomes large, the number of possible sequences increases exponentially, and the specific rate of exponential increase can be expressed as a simple function of p. This rate of increase, expressed on a per symbol basis, is denoted *entropy*, and takes the form

$$\begin{aligned} H(p) &\equiv \lim_{N\to\infty} \log \binom{N}{pN} / N \\ &= -p\log(p) - (1-p)\log(1-p). \end{aligned} \tag{1}$$

The simplified form of $H(p)$ for large N comes about as a direct result of Stirlings formula (Goldman, Appendix 2). In concrete terms, Equation 1 says that when N is large, there are about $2^{N\cdot H(p)}$ binary sequences of length N consisting of pN copies of the symbol A and $(1-p)N$ copies of the symbol B.

The total number of binary sequences of length N is 2^N, and binary sequences with symbols of probability p and $1-p$, and with length N large, are characterized by $0 \leq H(p) \leq 1$. Thus the ratio of binary sequences characterized by a given symbol probability p to all possible binary sequences is $2^{N\cdot H(p)-N}$. When $p < 1/2$, this ratio can be made arbitrarily small by increasing N. Also important is that all $2^{N\cdot H(p)}$ binary sequences with symbol probability p are equally likely. This makes sense because each sequence is composed of the same numbers of each type of symbol: only the arrangements of symbols differ.

We can use ideas similar to those associated with Equation 1 to characterize channel behaviors. Informally, this comes about because a given channel input can, due to random transitions, or "noise," generate any one of many different outputs, and we can count the number of inputs that could have caused a given output. A channel that is good for communication will have very few input sequences implied by a given output, while a poor channel will allow almost any input to generate any output.

More specifically, for a long received sequence at the output of Figure 1, each received $\tilde{B}$ could have been caused by transmission of either an A or B. The conditional probabilities $p(A|\tilde{A})$ and $p(B|\tilde{A}) = 1 - p(A|\tilde{A})$ can be used in Equation 1 to count the input sequences associated with the $\tilde{A}$'s, and similar computations can, in general, be carried out for the $\tilde{B}$'s. An average of these counts, weighted according to the relative frequencies of $\tilde{A}$'s and $\tilde{B}$'s, yields a count of the inputs that could have caused a given output.

An expression of this type is denoted $H(X|Y)$, and is called the *conditional entropy* or the *equivocation* of a channel. The conditional entropy $H(X|Y)$ of a channel is closely related to the *capacity* of the channel, and is a function of only the input symbol probabilities and the transition probabilities. As with unconditioned entropies, we can say (roughly) that for large N, an input of N symbols can generate about $2^{N\cdot H(Y|X)}$ output sequences, and an output of N symbols could have been caused by about $2^{N\cdot H(X|Y)}$ input sequences.

This brief sketch of communication concepts can be summarized as follows. A source can be characterized by the number of long sequences of length N that it can generate, subject to the probabilities that define the source. Likewise, a channel can be characterized by the number of output sequences it can generate for a given input of N symbols. Both of these quantities grow exponentially with N, and the entropy (for sources), or conditional entropy (for channels), specifies the specific rate of exponential growth with N.

III. A Deception Model Based on the Z-channel

Our analysis technique for deception is fundamentally based on the following definition. The term *deception target*, or just *target*, refers to anyone who has a deception applied against them, whether or not they "fall" for it. This definition is based on page 2 of [2].

Definition 1: Deception is the presentation of a specific false version of reality by a deceiver to a target for the purpose of changing the targets actions in a specific way that benefits the deceiver.

That is, deception is the imposition of a *specific* false version of reality onto an adversary: a deceiver does not simply cloak reality in an obscuring fog, but rather replaces reality with a specific and carefully created false version. Deception is thus quite distinct from the denial of information to an adversary, and it is quite distinct from efforts that direct an adversary in a random, haphazard direction. As stated eloquently in [15], a successful deception will make an adversary "...quite certain, very decisive, *and wrong*" [emphasis in the original].

This definition can be made precise. If an environment can be in any one of several states, and if agents within the environment use sensory data to infer the state of the environment, then the a deceivers actions consist abstractly of manipulating sensory data such that observations made by the target will suggest a specific incorrect version of reality.

These ideas can be interpreted in terms of Figure 1. The possible values for a state variable of interest are denoted on the left by A and B, and the inferences by the deception target about that value are denoted on the right by $\tilde{A}$ and $\tilde{B}$. Under normal circumstances, a target will be able to correctly infer the state of nature using available sensory data; for simplicity, we ignore ordinary "honest" mistakes. The actions of a deceiver cause a targets inferences about values of state variables to differ from the true values. These actions are modeled by a transition probability p_{BA}, which is the fraction of successful deceptions out of all deception attempts. We assume that this probability is measurable and known to both deceivers and deception targets.

A. Examples and Further Discussion

As a model for deception, the Z-channel of Figure 1 is, we believe, as general and valid a description of deception as Definition 1. To illustrate how our model can be used to describe a specific deception, consider a salesperson with a large number of items to sell. Some of the items happen to be of low quality, and the salesperson deceptively portrays these low quality items as high quality items. Here the state variable of interest is the quality of the item being sold in any given sales encounter, and this state variable can take on the values "High Quality," denoted B, and "Low Quality," denoted A, with probabilities p_A and $p_B = 1 - p_A$ respectively. As with transition probabilities, we assume that these "symbol" or "input" probabilities are measurable and thus known to both targets and deceivers. We also presume that if the salesperson were absent, or not deceptive, a purchaser would correctly determine a given items quality. That is, when no deception attempts are made the target observes the environment in a noiseless, or error-free, way.

However, a deceptive salesperson introduces noise into this sensor, with successful deceptions resulting in some number of low quality items appearing, to a purchaser, as high quality. The relative frequency at which this deception succeeds–that is, the relative frequency at which low quality items appear as high quality–is given by the transition probability p_{BA}.

In order for this deception to be attempted, it is necessary that the state variable hold the value A. That is, the salesperson can only attempt to make a low quality item appear as high quality when "pitching" a low quality item; the deception cannot be applied to a customer who is considering the purchase of what happens to be a high quality item. Because the deceivers actions can only cause one type of error, deceptions of this type–that is, those that can be modeled as a Z-channel–are referred to as *one sided* deceptions. We denote the value A as the *precipitating* value for the deception (though it might also be termed the *actual* value). The value A is a sort of *precondition* necessary for launch of the deception. Value B is the false version of reality referred to in Definition 1. We refer to B as the "false" or *bogus* value that the deceiver uses to mask, or disguise, value A.

The term "bogus" is just a label for a specific value of a state variable, and this label has meaning only with respect to a particular precipitating value in a particular deception. Of course, a state variable may actually take on the value B, an example being, again, a customer who happens to be considering a high quality item for purchase. More generally, the state variable in question may take on a value unrelated to the deception; that is, the state variable may take on a value that is distinct from the precondition for a given deception, and distinct from the bogus value for that deception. In our "deceptive salesperson" example, every item for sale is either high or low quality; we allow no other possibilities. For any specific deception we will, for simplicity, consider only the specific version of reality that acts as a pre-condition for the deception, and the false version of reality associated with the deception. In effect we are conditioning all probabilities on the event $A \vee B$. This assumption allows us to treat deception as a binary channel, as opposed to an m-ary channel in which only a single transmitted symbol is subject to noise. It is because the precipitating and false versions of reality are mutually exclusive, and because we are interested only in cases where one or the other hold, that the mathematical tools developed for binary communication system can be applied to deception.

Income tax evasion is another deception that can be described in terms of a Z-channel. In this case, a deceptive taxpayer presents to the taxing agency evidence (in the form of false statements, false documents, etc.) for an incorrectly low value of income. The state variable of interest is the income of the taxpayer, which can take on the value A, "High," or B, "Low." The transition probability p_{AB} is the rate at which high income tax payers successfully portray themselves as low income. The tax agent who examines the claim must

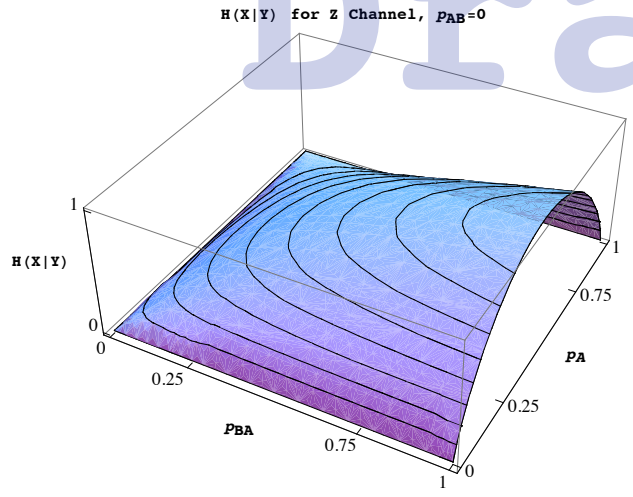

Fig. 3. Conditional entropy for Z-channel as a function of transition probability p_{AB} and input probability $p_A = 1 - p_B$. Contours follow constant $H(X|Y)$ values. This surface is not symmetric about $p_A = 1/2$.

evaluate all the available evidence (both evidence provided by the taxpayer, and that available from other sources) to determine the actual value of income. Only a high income taxpayer is capable of attempting this deception.

As with income tax evasion and sales of low quality items, camouflage can be modeled once we determine the precipitating value A and the bogus value B. For camouflage, some spatial region is being monitored by a sentry, and the task of the sentry is to determine whether an unauthorized person is present, a value that can be denoted A, or no unauthorized person is present, B. This deception can only be carried out by someone who is in the spatial region.

As a final example of a one-sided deception, consider identity theft. Many cases of social engineering [5] involve a deceiver who assumes the identity of an "insider." The state variable of interest here can take on either the value "this person has authority" which we'll denote B, or alternately "this person does not have authority," denoted A. The transition probability p_{AB} then represents the rate at which an unauthorized person is successful at being accepted as a person with authority.

IV. Conditional Entropy as a Measure of Deception Performance

The Z-channel model allows us to evaluate deception scenarios in the same ways we evaluate communication systems. Though communication channels are often characterized in terms of mutual information or capacity, we will use the closely related concept of conditional entropy because we believe it to more directly describe deception. In a given deception scenario, the conditional entropy $H(X|Y)$ indicates the number of possible sequences of values of the state variable of interest that could have caused the sequence of inferences made by the target. To simplify terminology, we will refer to a "sequence of values" of a state variable as an *SOV*, and we will refer to a "sequence of inferences" as an *SOI*. Thus, an SOV is associated with the environment, and an SOI is associated with the deception target. In a rough sense, the deception target wants and expects its SOI to the SOV, but the deceiver attempts to cause the SOI to systematically differ from the SOV.

When $p_{AB} = 0$ in Figure 1, we have $H(X|Y) = 0$, and there is either no deceiver, or the deceiver is completely ineffective. In this case, there is $2^{N \cdot 0} = 1$ possible SOV's that could have caused any particular SOI. Similarly, when $p_A = 0$ or $p_A = 1$, there is no uncertainty about the value of the state variable–the state variable never changes–and so again there is only one SOV, no matter what the observations imply.

A deceiver who is always successful is characterized by $p_{AB} = 1$ and $H(X|Y) = 1$. In this case, the deception target always decides that the bogus state of nature, state B, is in effect. Since the SOV and the SOI are completely independent, the number of SOV's that could have caused any particular SOI depends only on the probability p_A.

The Z-channel model of deception and the conditional entropy associated with it open many opportunities for analysis. The number of SOV's that could have caused a particular SOI increases exponentially with the number of deception attempts, and the conditional entropy is the rate of this exponential growth. Even though the number of SOV's increases exponentially with the number of deception attempts, when $H(X|Y) < 1$ the number of SOV's increases more slowly than the total number of possible sequences. Thus, as the number of deception attempts increases, the number of SOV's associated with a given SOI becomes an arbitrarily small fraction of the total number of possible sequences.

Our Z-channel model implies that the actual sequence of inferences realized in a long sequence of deception attempts is an arbitrary member of the class of possible SOI's. That is, the actual SOI realized by the deception target (or the deception target community) is no more probable and no less probable than any other sequence of inferences associated with a given set of input and transition probabilities. The actual SOI is not distinguished in any way from any other possible SOI.

A given value of conditional entropy can be achieved by many combinations of p_A and p_{AB}. As shown by the contours in Figure 3, a fixed value of $H(X|Y)$ can be maintained in spite of changes in p_A if compensating changes to p_{AB} can be made. In a similar way changes in deception performance due to changes in p_A and/or p_{AB} can be evaluated, as can cost effectiveness of proposed changes to, say, p_{AB}.

This model allows analysis of *compound deceptions*, which are multiple sub-deceptions coupled together in series or parallel. It also provides insight into "averages" of deceptions. Averages of deceptions are interesting because conditional entropy is a concave function of the input probabilities. Consider two distinct one-sided deceptions with transition probabilities $p_{1,AB}$ and $p_{2,AB}$, and conditional entropies $H_1(X|Y)$ and $H_2(X|Y)$. In regions where conditional entropy is a concave function of the transition probabilities, the average conditional

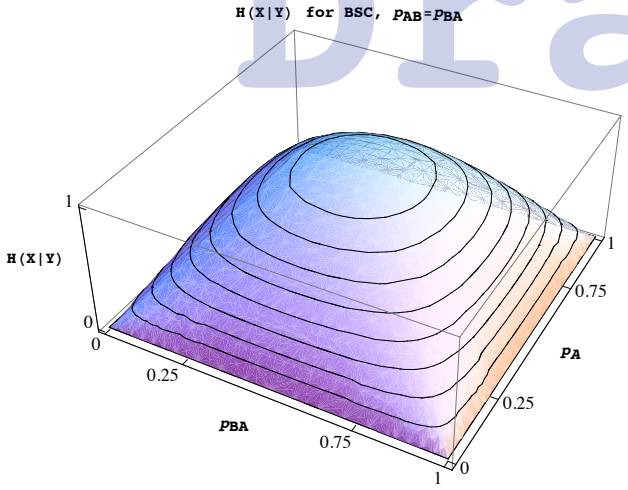

Fig. 4. Conditional entropy for a Binary Symmetric Channel as a function of transition probability $p_{AB} = p_{BA}$ and input probability $p_A = 1 - p_B$. Contours follow constant $H(X|Y)$ values. Because $p_A = 1 - p_B$, the conditional entropy shown here is for a special case of the channel shown in Figure 2

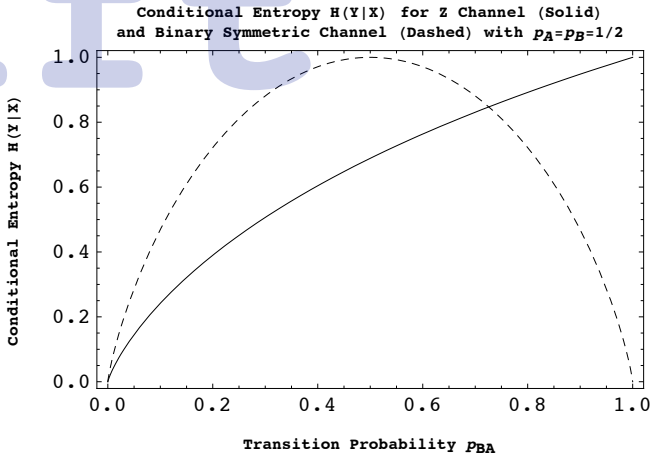

Fig. 5. Conditional entropy for Z-channel and Binary Symmetric Channel with $p_A = 1 - p_B = 1/2$. These curves are essentially "slices" through the surfaces in Figures 3 and 4.

entropy $(H_1 + H_2)/2$ will be less than the conditional entropy associated with $p_{Ave} \equiv (p_{1,AB} + p_{2,AB})/2$. This implies that a community of deception targets should prefer to distribute themselves uniformly between two deceivers, rather than commit themselves to a single deceiver characterized by p_{Ave}.

V. Symmetric Complements and Symmetric Deceptions

Although only one deception is being explicitly modeled in Figure 1, there are *two* distinct deceptions possible with the values A and B shown. As the figure stands, the deceiver causes value B to appear when value A is actually in effect. However, another possible deception is to cause value A to appear to hold when value B actually does. In other words, Figure 1 portrays only one of the two possible Z-channels that can be constructed between a binary input and binary output.

For most deceptions, the associated *symmetric complement* does not support, and may actually counter, the deceivers interests. For example, the deceptive portrayal of a low quality item as high quality has, as a symmetric complement, the portrayal of a high quality item as low quality, and the symmetric complement of income tax evasion is the depiction of a low income individual as one of high income.

However, there are some conditions under which a deception and its symmetric complement can together support the interests of a deceiver better than either deception alone. An example occurs in the following exchange over the the use of "dummy" aircraft to divert attacks away from real aircraft [14].

> Sometime around mid-1942, Major Oliver Thynne was a novice planner with Colonel Dudley Clarke's "A" Force, the Cairo-based British deception team. From intelligence, Thynne had just discovered that the Germans had learned to distinguish the dummy British aircraft from the real ones because the flimsy dummies were supported by struts under their wings. When Major Thynne reported this to his boss, Brigadier Clarke, the "master of deception," fired back
>
> "Well, what have you done about it?"
>
> "Done about it, Dudley? What could I do about it?"
>
> "Tell them to put struts under the wings of all the real ones, of course!"

Here dummy aircraft are used to deceive enemy attackers about the numbers and locations of real aircraft, with the state variable being items on a runway, with possible values "real aircraft" and "dummy aircraft." The only error possible for the deception targets (the attackers) is to mistakenly believe that a dummy aircraft is a real aircraft; the deception is thus one-sided.

However, this one-sided deception is not very effective: it is impractical for the British defenders to increase p_{AB} to an acceptable level. However, sheer intuition suggests that the original deception be combined with its symmetric complement. The one-sided deception is imperfect, and so the attackers tend not to make the only mistake they *can* make under this deception: namely, the attackers tend not to believe that the dummy aircraft are real. However, the symmetric complement makes the attackers vulnerable to two types of mistakes: dummy aircraft can be mistaken for real, and real aircraft can be mistaken for dummy.

In those cases where a deception along with its symmetric complement "make sense," the two acting together can potentially provide a deceiver with increased performance and flexibility. This can be seen by comparing Figure 3, the conditional entropy associated with a one-sided deception, with Figure 4, the conditional entropy associated with a BSC with $p_{AB} = p_{BA}$. (Note that Figure 4 is the conditional entropy for a special case of the channel shown in Figure 2.) The deceiver wants $H(X|Y)$ to be as large as possible. When H is one, slope is zero for BSC, but has a non-zero value for Z Channel. This means that for a symmetric deception,

want to be successful "about" half the time; for a one-sided deception, need to be successful all the time, or performance drops off. That is, a one-sided deception is "fragile" compared to a symmetric deception. This is illustrated in Figure 5.

VI. Deception for Computer Security

The analysis methods described above can be applied to two deception-based techniques for computer security that we are developing at the Naval Postgraduate School. These two techniques are *False Honeypots*, which can be used as a sort of "inoculation" against computer intrusions, and the *Spoofing Channel* which provides a method for responding to a successful intrusion once it has been detected.

A. Honeypots and False Honeypots

One of the most effective ways of gathering information about computer intruders is through the use of *honeypots* [3], which are computers placed on a network for the sole purpose of being broken into. Honeypots contain no information of value, and are usually highly instrumented so that the maximum amount of information about intruders and their activities can be gathered. Most computer intruders avoid honeypots to protect their intrusion techniques, which usually require significant expertise to develop. Honeypots are often "tricked out" to appear as a non-honeypot computer containing valuable information.

A honeypot thus deceptively appears as an ordinary computer, and in Figure 1 B can be assigned the value "this computer is ordinary," and A can be assigned the value "this computer is a honeypot." The symmetric complement of this deception is to portray an ordinary computer as a honeypot. Deployment of this symmetric complement has the potential to reduce or eliminate the value of a computer intruders observations about whether a successfully compromised computer is an ordinary computer or honeypot. That is, the intruders observations about the machine they have broken into, gathered by examining all different aspects of the machine, would provide little or no useful guidance for determining whether the machine is a honeypot or an ordinary computer.

B. Spoofing Channels

A spoofing channel is similar to an ordinary communication channel: both accept inputs consisting of sequences of symbols, and both deliver outputs consisting of strings of symbols. These channels differ in how their outputs are related to their inputs. The output of a communication channel allows the input to be approximated, and ideally, to be perfectly reconstructed. In essence, the input to a communication channel can be thought of as a "choice" of some member of a set, and the output aids, to a greater or lesser extent, in determining that choice.

In contrast, the output of a spoofing channel is only required to have the same statistical structure as the input. That is, the output of a spoofing channel is only required to have the same relative frequency of individual symbols, of symbol pairs, of symbol triplets, of symbol quadruplets, and so on, as the input. There need not be any stronger relationship between the input and the output of a spoofing channel.

Functionally, a spoofing channel is like a communication channel in that both resolve uncertainty, but a spoofing channel resolves uncertainty in the wrong way. Alternately, a spoofing channel can be thought of as a communication system that is broken, but such that the user at the receiving end cannot tell that it is broken.

We believe the spoofing channel to be a valuable deception-based information security tool, with perhaps its most obvious application as an Intrusion Response System (IRS). Whenever a computer Intrusion Detection System (IDS) such as SNORT [12] detects an in-progress intrusion, the system administrator responsible for the compromised machine must choose one of two unappealing courses of action. One option is to immediately drop the connection to the intruder, which ensures that sensitive information is protected to the maximum extent possible. This option has the significant disadvantage that the intruder is unequivocally notified that they have been detected. It would be much better if the connection could be maintained and the intruders activities on the target machine observed. This could aid forensic work and provide guidance for preventing future intrusions. However, this option is unacceptable because it leaves sensitive data vulnerable.

An IRS based on the spoofing channel ensures that an intruder never receives an original document from the compromised computer. Instead, an intruder would be able to retrieve or manipulate only spoofs, or impersonations, of original documents. The material made available to the intruder would have the same statistical structure as the original, but no stronger relationship. Sensitive material would thus be protected, but it would be difficult or impossible for an intruder to tell that it was bogus. A spoofing channel does not have to work perfectly to be useful as an intrusion response system; forcing an intruder to spend time and effort to determine the validity of collected data can be counted as a success.

Close relatives of the spoofing channel have been observed "in the wild." Examples include the classic spoof created by hand by Alan Sokal [13], the *SCIgen* software for automatically generating random computer science research papers [9], and the ELIZA program. These examples are functionally equivalent to spoofing channels that have built-in repositories of non-spoofed data.

The most significant characteristic of the spoofing channel is that it exploits for deception the uncertainty that conventional communication resolves. The output of a channel by itself provides no direct evidence for deciding whether that channel is communicating or spoofing. That is, a channels output is not a useful observation for deciding whether that channel is delivering the associated input, or a statistically valid spoof of that input. The only tool available at the output of a channel for deciding whether to treat it as a communication or a spoofing channel is the prior probabilities.

Our work so far has focused on spoofing channels for natural language text. Two fundamental techniques have suggested

TABLE I
SUMMARY OF DECEPTION EXAMPLES.

Example	Deceiver	Target	State Variable X	(Precipitating, Bogus) Values	Deception	Symmetric Complement
Sale of Shoddy Item	Seller	Purchaser	Quality of Item	Item is (Shoddy, High Quality)	Shoddy Item Promoted as High Qual.	High Quality Item Promoted as Shoddy
Tax Evasion	Taxpayer	Tax Agent	Income of Taxpayer	Income is (High, Low)	High Income Taxpayer Posing as Low Income	Low Income Taxpayer Posing as High Income
Camouflage	Intruder	Sentry	Location of Intruder	Intruder is (In, Not In) Region	Intruder Appears Not In Region when Is	Intruder Appears to be In Region When Not
Identity Theft	Person w/o Authority	Functionary	Identity of Claimant	Claimant (Is Not, Is) Authorized	Person w/o Authority Claims Authority	Person w/Authority Professes None
Runway Strafing	Runway Defenders	Attacking Pilots	Status of Runway	Runway (Is, Is Not) Usable	Usable Runway Disguised as Unusable	Unusable Runway Disguised as Usable
Honeypot	Computer Administrator	Computer Intruder	Type of Computer	Computer is (Honeypot, Ordinary)	Honeypot Disguised as Ordinary Computer	Ordinary Computer Disguised as Honeypot
Fake Honeypot	Computer Administrator	Computer Intruder	Type of Computer	Computer is (Ordinary, Honeypot)	Ordinary Computer Disguised as Honeypot	Honeypot Disguised as as Ordinary Computer
Spoofing Channel	Computer Administrator	Computer Intruder	Status of Intrusion	Intrusion (has, has not) been Detected	Bogus Information Delivered as Valid	Valid Information Delivered as Bogus

themselves for automatic generation of spoofs of natural language text documents. One technique for modifying a documents meaning while maintaining its "style" structure is through manipulation of the target documents semantic structure. This is intuitively the most straightforward approach to changing the meaning of a document while maintaining the same "style." An example of this sort of technique would consist of negating and un-negating some particular subset of assertions in the subject document.

Another technique for automatically changing a documents meaning is through manipulations based on syntactic structure. A technique of this sort might consist of simply swapping two successive noun phrases (which may appear in the same, or in different, sentences). This technique depends heavily on *pareidolia*, which is the psychological phenomena of finding meaning in random and presumably ambiguous patterns [6].

VII. CONCLUSION

In this paper we have shown how the existing theory of communication can be used, almost "as is," to describe deception. Based on a natural and general definition of deception, our model clarifies the previously obscure relationship between deception and communication, and establishes that every deception has a symmetric complement. Ordinary (i.e., one-sided) deceptions can, when always successful, cause observations made by a deception target to become worthless; the observations become no better than the flip of a coin in guiding the targets activities. However, this situation is fragile in the sense that even isolated failures on the part of the deceiver are almost inevitably

In those cases where a deception and its symmetric complement can be sensibly deployed together, a targets observations of the environment can be made worthless in a stable, or robust, way: as above, the observations of a target can be made no more valuable than information gained by flipping a coin, but in this case isolated failures by the deceiver may be difficult for the target to exploit. However, if a deceiver moves beyond the point at which observations become worthless to a point where there is a systematic correlation between the false version of reality and the actual version of reality, the deceiver becomes vulnerable to counter-deception techniques. This situation is analogous to the minimax solution of a game.

Even more interesting than the specifics presented in this paper are the many open questions that remain. A few outstanding topics include the relationship of rate distortion theory [1] to deception; analysis of deception as a game [4] with payoffs quantified by mutual information; models of deception using continuous state variables; and the influence of deception on the stability of signaling systems [10].

REFERENCES

[1] Cover, Thomas M., and Joy A.Thomas. *Elements of Information Theory, 2nd Ed.* John Wiley and Sons, 2006.

[2] Godson, Roy, and James J. Wirtz, Eds. *Strategic Denial and Deception: The Twenty-First Century Challenge.* Transaction Publishers, New Brunswick, New Jersey, 2002.

[3] The Honeynet Project. *Know Your Enemy: Learning About Security Threats, 2nd Ed.* Addison-Wesley, New York, 2004.

[4] Garg, Nandan, and Daniel Grosu. *Deception in Honeynets: A Game-Theoretic Analysis.* Proceedings of the 2007 IEEE Workshop on Information Assurance, United States Military Academy, West Point, NY, 20-22 June 2007.

[5] Mitnick, Kevin D., and William L. Simon. *The Art of Deception: Controlling the Human Element of Security.* Wiley Publishing, Inc. Indianapolis, Indiana, 2002.

[6] http://en.wikipedia.org/wiki/Pareidolia

[7] Rowe, Neil C., Binh T. Duong, and E. John Custy. "Fake Honeypots: A Defensive Tactic for Cyberspace." Proceedings of the 7th IEEE Workshop on Information Assurance, U.S. Military Academy, West Point, New York, June 2006.

[8] Rowe, Neil C., Han C. Goh, Sze L. Lim, and Binh T. Duong. "Experiments with a Testbed for Automated Defensive deception Planning for Cyber-Attacks."

[9] http://pdos.csail.mit.edu/scigen/

[10] Searcy, William A., and Stephen Nowicki. *The Evolution of Animal Communication: Reliability and Deception in Signaling Systems.* Princeton University Press, Princeton, New Jersey, 2005.

[11] Shannon, Claude E., and Warren Weaver. *The Mathematical Theory of Information.* University of Illinois Press, Chicago, Illinois, 1949.

[12] http://www.snort.org

[13] http://www.physics.nyu.edu/faculty/sokal/

[14] Whaley, Barton. *Conditions Making for Success and Failure of Denial and Deception: Authoritarian and Transition Regimes*. Printed as Chapter 3 of [2].

[15] Whaley, Barton. *Stratagem: Deception and Surprise in War.* Artech House, Boston, Massachusetts, 2007.